Facilitating Multiple Target Tracking using Semantic Depth of Field (SDOF)

Nivedita R. Kadaba
University of Manitoba
Winnipeg, Manitoba
Canada
nrkadaba@cs.umanitoba.ca

Xing-Dong Yang
University of Alberta
Edmonton, Alberta
Canada
xingdong@cs.ualberta.ca

Pourang P. Irani
University of Manitoba
Winnipeg, Manitoba
Canada
irani@cs.umanitoba.ca

CHI 2009, April 4 – 9, 2009, Boston, MA, USA
ACM 978-1-60558-247-4/09/04.

Abstract
Users of radar control systems and monitoring applications have to constantly extract essential information from dynamic scenes. In these environments a critical and elemental task consists of tracking multiple targets that are moving simultaneously. However, focusing on multiple moving targets is not trivial as it is very easy to lose continuity, particularly when the objects are situated within a very dense or cluttered background. While focus+context displays have been developed to improve users' ability to attend to important visual information, such techniques have not been applied to the visualization of moving objects. In this paper we evaluate the effectiveness of a focus+context technique, referred to as Semantic Depth of Field (SDOF), to the task of facilitating multiple target tracking. Results of our studies show an inclination for better performance with SDOF techniques, especially in low contrast scenarios.

Authors Keywords
Semantic depth of field, moving targets, visual displays, visualization, blurring, preattentive cues, target tracking.

ACM Classification Keywords
H.5.2 [Information interfaces and presentation (e.g. HCI)]: User Interfaces---Graphical User Interfaces, Evaluation/Methodology, User-centered design, Interaction styles.

General Terms
Experimentation, Human Factors, Performance.

Introduction

A significant amount of information used in the information sciences is represented using dynamic simulations and animations. These include the display of traffic control systems, video games, and interactive maps. The effectiveness of these systems depends upon techniques that facilitate viewing visual scenes that are dynamically updated.

With dynamic information, a significant concern is to track all the changes that occur simultaneously in the scene. For example, in an interactive map of a city in a car navigation system, the user may be interested in isolating several objects of interest, such as hotels. Although the maps display ample information, they do not facilitate the isolation of items of interest. As a result, the task of finding necessary information can be quite overwhelming and time-consuming (Figure 1). This is especially true when there exists varying levels of contrast between the targets and the background. It is therefore important to devise techniques that allow users to isolate and focus on elements of the display that are deemed important at any given time.

In this paper, we examine the effects of applying a focus+context technique, semantic-depth-of-field (SDOF), to allow users to visually parse dense visual scenes. SDOF exploits the preattentive capabilities of the human eye such that elements of interest pop-out to the user. While SDOF was shown to be effective for preattentively processing information in static environments [3, 4], to the best of our knowledge there has not been any study investigating the benefits of this technique in dynamic scenes. Our primary contribution is the demonstration that SDOF can be an effective technique for tracking moving targets.

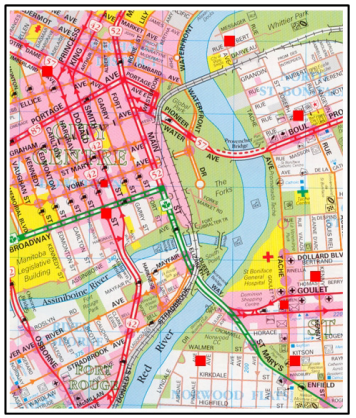

Figure 1. Look quickly! How many red targets can you pick out from the scene? This task requires significant visual resources to carry out. (appears better in color)

Related Work

Two areas of research specifically relate to this work: Multiple Target Tracking (MTT) and Semantic Depth of Field (SDOF).

Multiple Target Tracking (MTT)
There exist several contending theories that explain the mechanisms that our perceptual system employs for tracking multiple moving targets. Pylyshyn et al [6]

suggested that human beings simultaneously track multiple moving objects (targets) using a primitive mechanism, called the FINST (Fingers of Instantiation) model. The key idea of the FINST model is that "sticky" indexes can be assigned to the targets, which remain with the target even if it continuously changes its location. Pylyshyn et al contended that the human eye distinguishes targets in two stages; in the first stage the indexes are serially assigned to the targets (or in parallel if the targets "pop-out" [5]), and in the second stage the targets are tracked in parallel. Experiments showed that our visual system can track up to 4 or 5 targets simultaneously. Performance reduces significantly when the number of targets increases above this limit.

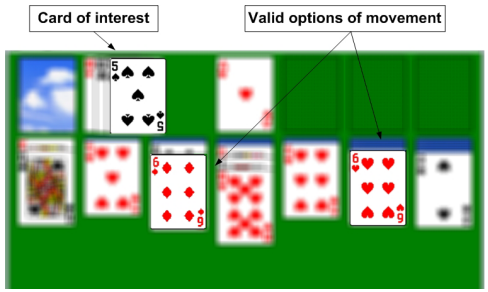

Figure 2. SDOF can be employed to highlight objects of interest, such as all valid movement options for a given card in a game of Solitaire (redrawn from Microsoft® Solitaire on Windows XP)

Cavanagh and Alvarez's [1] multifocal attention model suggests that each target receives a single attention channel that selects and tracks multiple targets in parallel. Yantis [8] suggested that targets are grouped into rigid polygons, which helps the participants track the targets according to their relative positions the scene. Using this method, participants were able to track up to a maximum of 5 targets. However, this efficiency depended directly on the rigidity of the polygon, and decreased drastically as a target violated the boundaries of this polygon.

Semantic Depth of Field (SDOF)

Visual overload is caused when users have difficulty distinguishing between important and unimportant events in a given scenario. Classic highlighting styles like color-based highlighting [2] direct user's attention to a particular area of view without distorting the context. Other techniques such as fish-eye views [7] are distortion-oriented as they enlarge the object of interest to catch the user's attention, and shrink the surrounding context in order to fit the scene in the limited size of the view-port.

Unlike previous highlighting techniques, SDOF [3] preserves the physical properties of the highlighted object, e.g. size, shape, color, and texture, while "distorting" (blurring) the objects of lesser interest. SDOF has several favourable characteristics (Figure 2):

- SDOF does not affect the physical properties of the objects.
- SDOF reduces the sharpness of objects that are unimportant and does not modify the spatial properties of critical objects.
- SDOF is useful in giving an overview of the information space and to draw attention to specific sections of the display.
- SDOF is intuitive, and unlike other types of focus+context techniques, does not require training.

In a series of experiments, Kosara [3, 4] showed that SDOF is preattentive and can be perceived by the human brain without serial search. Kosara's other findings suggest that SDOF should not be used as a visualization dimension because the participants are not able to tell the difference between objects with different blur levels in any meaningful ways. He also stated that SDOF did not work well with pixel-based visualizations. Although SDOF has shown promising applications, it has not been applied to the visualization of animated or dynamic objects.

Experiment

The goal of this experiment was to test the effect of SDOF in successfully isolating given targets in a visually dense scenario. Based on prior literature on SDOF and MTT, we formulated the following hypotheses:

- **Hypothesis 1:** Participants will perform target recognition more accurately using SDOF.
- **Hypothesis 2:** SDOF improves performance as the contrast weakens.

Participants

Seven students, between the ages of 20 – 30 years, from a local university participated in this experiment.

Materials

Our experiment comprised of a C# .NET program containing interactive maps and animated objects executed on a Windows XP machine, and displayed on a 17" monitor with a 1024x768 screen resolution.

Experimental Conditions

INDEPENDENT VARIABLES

The following variables were manipulated during the experiment:

- **Contrast:** Three levels of contrast between the background and the objects were employed; low, medium, and high.
- **SDOF level:** Two levels of SDOF were tested in the experiment. The "Normal" level was the control for the experiment and did not employ SDOF. SDOF – 3 and SDOF – 9 depicted SDOF with less and more blurriness respectively. The numbers 3 and 9 denote the number of pixels (surrounding each pixel) that were employed in the blurring algorithm.
- **Number of targets:** Three numbers of targets (3, 4, and 5), each of 12 x 12 pixels in size were displayed.

DEPENDENT VARIABLES

The following parameters were recorded during the experiment:

- **Accuracy:** Participant response was recorded at the end of each trial; a correct response was given a score of 1, while an incorrect response received a score of 0. Participant scores were averaged over repeated trials.
- **Response time:** The time taken to give an answer was recorded in seconds and averaged over repeated trials.

Tasks

In an effort to test our hypotheses, our experiment followed the Multiple Object Tracking methodology of Plyshyn et al [6] and consisted of three tasks, as described below:

- **Isolation:** In this task, the participants were asked to isolate the targets in the scenario. The targets were randomly chosen and displayed by highlighting and flashing them for a few seconds at the beginning of the trial.

- **Tracking:** On completion of the *isolation* task all the objects in the scene began travelling about in random paths (Figure 1). The participants' task was to remember the *isolated* targets to track them as they moved.
- **Response:** When the *tracking* task was completed, one object in the scene was surrounded with a square box. The participant was asked to respond if the highlighted object was among the targets that were shown in the *isolation* task or not.

Procedure
The experiment was conducted in two phases. In the *demonstration* phase the participant asked to practice on a sample version of the system.

In the *experimental* phase the participant was shown several scenarios which encompassed a partial Latin-square of our experimental conditions (contrast, SDOF level, number of targets).At the beginning of this phase, a map was shown on the screen which contained the objects in the trial. The targets were displayed with black boxes and flashed 5 times for emphasis. After displaying the targets all the objects started moving inside the map in random paths. Objects were not allowed to occlude each other and were restricted to the area enclosed within the map. After 10 seconds, the objects stopped moving and a random object was surrounded with a black box. The participants were asked to respond (Y/N) whether they thought the highlighted object was a target or not.

Results and Discussion
We collected accuracy rates and response times from each trial and averaged this data across all participants. An analysis of our data showed that there were no main effects caused by the number of targets. We therefore combined these results and present them without differing to the target number. A one-way ANOVA did not reveal any significant differences in accuracy rates when SDOF was applied to enhance target tracking. The averages indicate that users were as accurate with SDOF as without in low, medium and high contrast conditions. We attribute this result to the limited number of participants for this experiment (Figure 3).

However, on average response times show a trend in favor of SDOF (Figure 4). Although this result was also not significant, we believe the trend partially supports our first hypothesis. We particularly notice that the difference between SDOF and the normal condition is greater in the low contrast condition. These response times indicate that SDOF may be more useful in the low contrast scenarios than in high contrast scenarios, which concurs with our second hypothesis. We believe that as the contrast in the scene reduced SDOF allowed users to better identify the targets.

Conclusion and Future Work

This study evaluates relatively a new technique that enhances visual processing in a highly cluttered scenario. SDOF has been previously used in static scenarios [3, 4] to enhance important events in the scene. This study is a first step aimed at testing the efficiency of SDOF in highly dynamic scenarios. In our study we tested the effectiveness of SDOF in enhancing a multiple target tracking task. Results of our study showed that participants were ~10% faster when SDOF was employed to highlight targets of interest in the scene (although this result was not significant). In addition, our results also showed that in low contrast scenarios, where it is more difficult to distinguish

between the background and the targets, scenes that were enhanced with the SDOF technique showed better response times when compared to scenes that were not enhanced with SDOF (although non-significant).

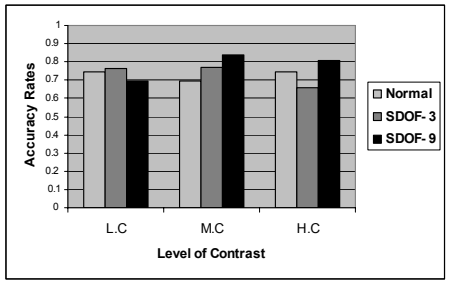

Figure 3. Accuracy rates for tracking targets in low, medium or high contrast scenarios

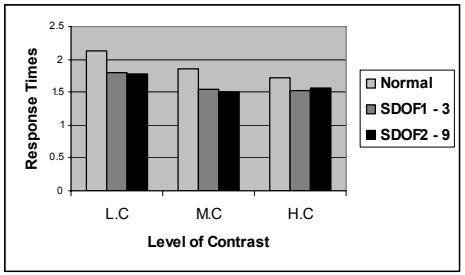

Figure 4. Response times for tracking 3, 4 or 5 targets in low, medium or high contrast scenarios

Lack of significance could have been due to the low number of participants in our experiment. We believe this would be different if we selected more participants to run the study. Overall, our results show promise toward employing SDOF to enhance scenarios that are highly dynamic, cluttered, and require extensive visual processing in order to sort out critical and non-critical events. Additional testing is however required to determine the degree to which SDOF improves the efficiency of multiple target tracking in a dynamic scene. Future work in this area also includes testing this technique in a practical scenario, such as a GPS system, in order to determine to what extent SDOF can be applied to everyday use and context.

References and Citations

[1] Cavanagh, P. and Alvarez, G. Tracking multiple targets with multifocal attention. Trends in Cognitive Science, 9(7), (2005), 249 – 354.

[2] Heer, J., Card, S.K., and Landay, J.A. Prefuse: A toolkit for interactive information visualization. Proceedings of the ACM SIGCHI conference on Human factors in Computing Systems, (2005), 421 – 430.

[3] Kosara, R., Miksch, S., and Hauser, H. Focus+Context taken literally. Proceedings of the IEEE Computer Graphics and Applications, 22(1), (2002), 22 – 29.

[4] Kosara, R., Miksch, S., Hauser, H., Schrammel, J., Giller, V., and Tscheligi, M. Useful properties of Semantic Depth of Field for better F+C visualization. Proceedings of the symposium on Data Visualization. Eurographics Association, (2002), 205 – 210.

[5] Pylyshyn, Z.W. and Annan, V. Dynamics of target selection in Multiple Object Tracking. Space Vision, 19(6), (2006), 485 – 504.

[6] Pylyshyn, Z.W. and Storm, R.W. Tracking multiple independent targets: Evidence for a parallel tracking mechanism. Spatial Vision, 3(3), (1998), 179 – 197.

[7] Sarkar, M. and Brown, M.H. Graphical fisheye views. Communications of the ACM, 37(12), (1994), 73 – 83.

[8] Yantis, S. Multielement Visual Tracking: Attention and Perceptual Organization. Cognitive Psychology, 24(3), (1992), 295 – 340.